

NEW WHEELS (SQL-Intro)

Apr B Sunday Onkar 10:30 AM Batch

Arindam Saha

1. Context

A lot of people in the world share a common desire: to own a vehicle. A car or an automobile is seen as an object that gives the freedom of mobility. Many now prefer pre-owned vehicles because they come at an affordable cost, but at the same time, they are also concerned about whether the after-sales service provided by the resale vendors is as good as the care you may get from the actual manufacturers. New-Wheels, a vehicle resale company, has launched an app with an end-to-end service from listing the vehicle on the platform to shipping it to the customer's location. This app also captures the overall after-sales feedback given by the customer.

2. Objective

New-Wheels sales have been dipping steadily in the past year, and due to the critical customer feedback and ratings online, there has been a drop in new customers every quarter, which is concerning to the business. The CEO of the company now wants a quarterly report with all the key metrics sent to him so he can assess the health of the business and make the necessary decisions.

As a data analyst, you see that there is an array of questions that are being asked at the leadership level that need to be answered using data. Import the dump file that contains various tables that are present in the database. Use the data to answer the questions posed and create a quarterly business report for the CEO.

3. Data Description

The data provided has

- Attributes on the vehicles New-Wheels sells - What are the make, model, and year? What is the price point?
- Attributes on the customers, such as where they live and payment methods
- Attributes on orders and shipments, such as when the order was shipped and received, what the after-sales feedback was, and so on.

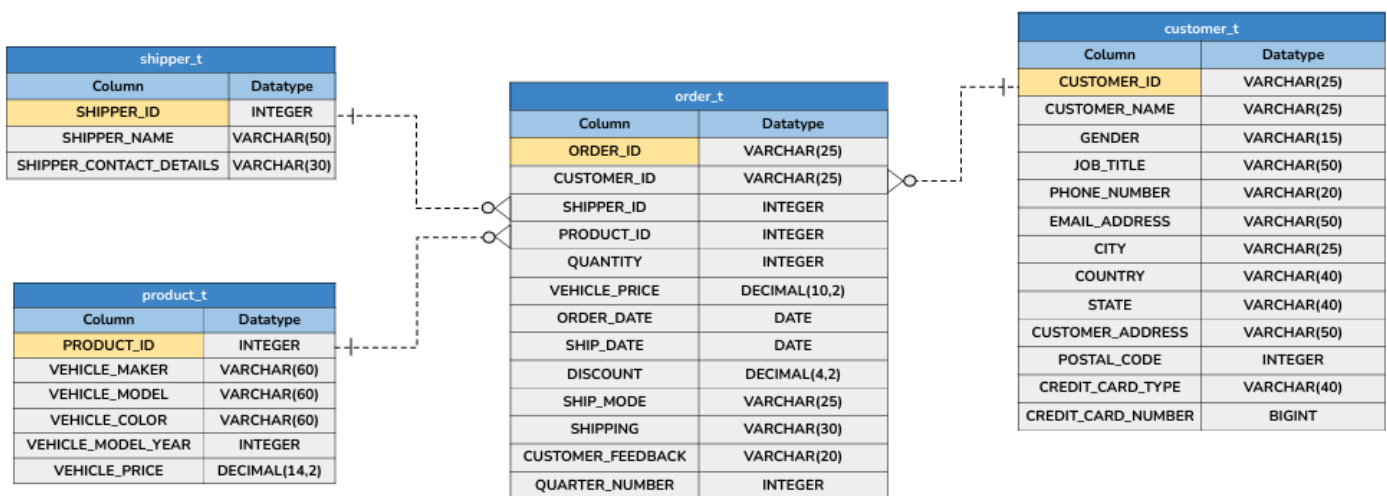


Figure 1: Entry Relationship Diagram

4. Business Questions:

4.1. Question 1:

Find the total number of customers who have placed orders. What is the distribution of the customers across states?

4.1.1. Solution Query:

```
select state, count(distinct customer_id) as Distribution_of_Customer from customer_t
where customer_id in (select distinct customer_id from order_t)
group by state order by Distribution_of_Customer desc;
```

4.1.2. Output:

Query 1 **Passed**: select state, count(distinct customer_id) as Distribution_of_Customer from customer_t where customer_id in (select distinct customer_id from order_t) group by state order by Distribution_of_Customer desc
Output:
Showing first 10 rows out of 37 rows

state	Distribution_of_Custo...
California	17
Texas	10
Florida	9
New York	7
Virginia	5
Michigan	5
Illinois	5
District of Columbia	5
Pennsylvania	4
Ohio	4

4.1.3. Observations:

- California (17 customers) has the highest number of customers followed by Texas (10 customers) and Florida (9 customers).
- Top 10 rows have been attached.

4.2. Question 2:

Which are the top 5 vehicle makers preferred by the customers?

4.2.1. Solution Query:

```
select PT.vehicle_maker, count(distinct OT.customer_id) as total_no from product_t PT
join order_t OT
on PT.product_id = OT.product_id
```

group by PT.vehicle_maker order by total_no desc limit 5;

4.2.2. Output:

Query 1 (Passed): select PT.vehicle_maker, count(distinct OT.customer_id)
desc limit 5

Output:

Showing 5 rows

vehicle_maker	total_no
Chevrolet	83
Ford	63
Toyota	52
Pontiac	50
Dodge	50

4.2.3. Observations:

- The top five makers are Chevrolet followed by Ford, Toyota, Pontiac, and Dodge.

4.3. Question 3:

Which is the most preferred vehicle maker in each state?

4.3.1. Solution Query:

```
SELECT state, vehicle_maker, total_customers, rank_no from (  
  SELECT PT.vehicle_maker,  
         CT.state,  
         sum(OT.customer_id) AS total_customers,  
         RANK() OVER (PARTITION BY CT.state ORDER BY sum(DISTINCT OT.customer_id) desc) AS rank_no  
  FROM product_t PT  
  JOIN order_t OT  
    ON PT.product_id = OT.product_id
```

```

JOIN customer_t CT
    ON OT.customer_id = CT.customer_id

GROUP BY CT.state, PT.vehicle_maker

Order by CT.state, PT.vehicle_maker
)

WHERE rank_no = 1

order by total_customers desc;

```

4.3.2. Output:

Query 1 (1/10/2020) : SELECT state, vehicle_maker, total_customers, rank_no from (SELECT PT.vehicle_maker, CT.state, sum(OT.customer_id) AS total_customers, RANK() OVER (PARTITION BY CT.state ORDER BY sum(DISTINCT OT.customer_id) desc) AS rank_no FROM product_t PT JOIN order_t OT ON PT.product_id = OT.product_id JOIN customer_t CT ON OT.customer_id = CT.customer_id GROUP BY CT.state, PT.vehicle_maker Order by CT.state, PT.vehicle_maker) WHERE rank_no = 1 order by total_customers desc

Output:

Showing first 10 rows out of 37 rows

state	vehicle_maker	total_customers	rank_no
California	Pontiac	893	1
Florida	Ford	880	1
Texas	Nissan	785	1
Indiana	Mazda	756	1
Tennessee	Mitsubishi	641	1
Iowa	Porsche	615	1
Nevada	Pontiac	615	1
Virginia	Volkswagen	615	1
New York	BMW	603	1
South Carolina	Isuzu	603	1

4.3.3. Observations:

- For Top selling state California, Pontiac is the top vehicle maker.
- Only 10 rows are fetched.

4.4. Question 4:

Find the overall average rating given by the customers. What is the average rating in each quarter?

4.4.1. Solution Query:

```

select quarter_number, avg(feedback_rating) as AVG_rating_per_q from (
select *,
case when customer_feedback = 'Very Bad' then 1
when customer_feedback = 'Bad' then 2
when customer_feedback = 'Okay' then 3

```

```

when customer_feedback = 'Good' then 4
when customer_feedback = 'Very Good' then 5
else 0 end as feedback_rating
from order_t)
group by quarter_number;

```

4.4.2. Output:

SQL queries passed

Query 1 **Succeeded**: select quarter_number, avg(feedback_rating) as AVG_rating_per_q from (select *, case when customer_feedback = 'Very Bad' then 1 when customer_feedback = 'Bad' then 2 when customer_feedback = 'Okay' then 3 when customer_feedback = 'Good' then 4 when customer_feedback = 'Very Good' then 5 else 0 end as feedback_rating from order_t) group by quarter_number

Output:

Showing 4 rows

quarter_number	AVG_rating_per_q
1	3.554838709677419
2	3.354961832061069
3	2.9563318777292578
4	2.3969849246231156

4.4.3. Observations:

- Avg rating for 1st quarter, 2nd quarter, 3rd quarter, and 4th quarter are 3.55, 3.35, 2.95, and 2.39 respectively.

4.5. Question 5:

Find the percentage distribution of feedback from the customers. Are customers getting more dissatisfied over time?

4.5.1. Solution Query:

```

SELECT
quarter_number,
SUM(CASE WHEN customer_feedback = 'Very Bad' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS
very_bad_percentage,
SUM(CASE WHEN customer_feedback = 'Bad' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS bad_percentage,
SUM(CASE WHEN customer_feedback = 'Okay' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS okay_percentage,
SUM(CASE WHEN customer_feedback = 'Good' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS good_percentage,
SUM(CASE WHEN customer_feedback = 'Very Good' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS
very_good_percentage
FROM
order_t
GROUP BY

```

quarter_number

ORDER BY

quarter_number;

4.5.2. Output:

SQL queries passed

Query 1 **(Passed)**: SELECT quarter_number, SUM(CASE WHEN customer_feedback = 'Very Bad' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS very_bad_percentage, SUM(CASE WHEN customer_feedback = 'Bad' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS bad_percentage, SUM(CASE WHEN customer_feedback = 'Okay' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS okay_percentage, SUM(CASE WHEN customer_feedback = 'Good' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS good_percentage, SUM(CASE WHEN customer_feedback = 'Very Good' THEN 1 ELSE 0 END) * 100.0 / COUNT(*) AS very_good_percentage FROM order_t GROUP BY quarter_number ORDER BY quarter_number

Output:

Showing 4 rows

quarter_number	very_bad_percentage	bad_percentage	okay_percentage	good_percentage	very_good_percentage
1	10.96774193548387	11.290322580645162	19.032258064516128	28.70967741935484	30
2	14.885496183206106	14.122137404580153	20.229007633587788	22.137404580152673	28.625954198473284
3	17.903930131004365	22.707423580786028	21.83406113537118	20.96069868995633	16.593886462882097
4	30.65326633165829	29.14572864321608	20.100502512562816	10.050251256281408	10.050251256281408

4.5.3. Observations:

- Quarter 1 has the highest positive feedback followed by 2nd quarter.
- Quarter 4 has the lowest negative feedback followed by 3rd quarter.

4.6. Question 6:

What is the trend of the number of orders by quarter?

4.6.1. Solution Query:

select quarter_number, count(distinct order_id) as count_orders from order_t group by quarter_number;

4.6.2. Output:

Query 1 **(Passed)**: select quarter_number, count(distinct order_id) as count_orders from order_t group by quarter_number

Output:

Showing 4 rows

quarter_number	count_orders
1	310
2	262
3	229
4	199

4.6.3. Observations:

- Quarter 1 has highest number of orders got placed where as in quarter 4 least number of orders were placed.

4.7. Question 7:

Calculate the net revenue generated by the company. What is the quarter-over-quarter % change in net revenue?

4.7.1. Solution Query:

select

quarter_number, a.revenue, lag(a.revenue,1,0) over(order by quarter_number) as Prev_q_revenue,

(a.revenue-(lag(a.revenue,1,0) over(order by quarter_number)))/(lag(a.revenue,1,0) over(order by quarter_number))*100
as QOQ

from

(

select quarter_number, SUM(QUANTITY*(VEHICLE_PRICE-(VEHICLE_PRICE*DISCOUNT))) as revenue

from order_t group by quarter_number

) as a group by quarter_number order by quarter_number ;

4.7.2. Output:

SQL queries passed

Query 1 **Passed**: select quarter_number, a.revenue, lag(a.revenue,1,0) over(order by quarter_number) as Prev_q_revenue, (a.revenue-(lag(a.revenue,1,0) over(order by quarter_number)))/(lag(a.revenue,1,0) over(order by quarter_number))*100 as QOQ from (select quarter_number, SUM(QUANTITY*(VEHICLE_PRICE-(VEHICLE_PRICE*DISCOUNT))) as revenue from order_t group by quarter_number) as a group by quarter_number order by quarter_number

Output:

Showing 4 rows

quarter_number	revenue	Prev_q_revenue	QOQ
1	18032549.899600014	0	
2	13122995.756199999	18032549.899600014	-27.22606714377602
3	8882298.844900003	13122995.756199999	-32.31500634522774
4	8573149.280599996	8882298.844900003	-3.48051298091048

4.7.3. Observations:

- QOQ has reduced over the quarters.
- Transition from 2nd to 3rd quarter has the highest impact of 32.3% drop in revenue. Then the transition from 1st to 2nd being the second at 27.2% drop in revenue.

4.8. Question 8:

What is the trend of net revenue and orders by quarters?

4.8.1. Solution Query:

select

```

quarter_number, SUM(QUANTITY*(VEHICLE_PRICE-(VEHICLE_PRICE*DISCOUNT))) as net_revenue,
COUNT(*) AS order_count

from order_t

group by quarter_number

order by quarter_number;

```

4.8.2. Output:

SQL queries passed

Query 1 (Passed): select quarter_number, SUM(QUANTITY*(VEHICLE_PRICE-(VEHICLE_PRICE*DISCOUNT))) as net_revenue, COUNT(*) AS order_count
 quarter_number
 Output:
 Showing 4 rows

quarter_number	net_revenue	order_count
1	18032549.899600014	310
2	13122995.756199999	262
3	8882298.844900003	229
4	8573149.280599996	199

4.8.3. Observations:

- In 1st quartile the net revenue being the highest (sum amount 18032549.9) against 310 number of orders.
- In last quartile the net revenue being the lowest (sum amount 8573149.28) against 199 number of orders.

4.9. Question 9:

-- What is the average discount offered for different types of credit cards?

4.9.1. Solution Query:

```

select credit_card_type, avg(discount) as AVG_discount, count(*) as No_of_customers

from customer_t CT

join order_t OT

on OT.customer_id = CT.customer_id group by credit_card_type order by AVG_discount desc;

```

4.9.2. Output:

SQL queries passed

Query 1 **Passed**: select credit_card_type, avg(discount) as AVG_discount, count(*) as No_of_customers from customer_t CT join order_t OT on OT.customer_id = CT.customer_id group by credit_card_type order by AVG_discount desc

Output:

Showing first 10 rows out of 16 rows

credit_card_type	AVG_discount	No_of_customers
instapayment	0.77	2
solo	0.7	1
americanexpress	0.6825000000000001	8
diners-club-enroute	0.665	6
diners-club-carte-blanc	0.6475000000000001	4
mastercard	0.64625	8
visa-electron	0.6380000000000001	5
maestro	0.6371428571428571	7

4.9.3. Observations:

- Highest discount offered to customer having instapayment credit card.
- Least discount offered to customer having diners-club-us-ca credit card.

4.10. Question 10:

What is the average time taken to ship the placed orders for each quarter?

4.10.1. Solution Query:

```
SELECT quarter_number, AVG(julianday(ship_date) - julianday(order_date)) AS avg_shipping_time
```

```
FROM order_t
```

```
GROUP BY quarter_number
```

```
ORDER BY quarter_number;
```

4.10.2. Output:

SQL queries passed

Query 1 **Passed**: SELECT quarter_number, AVG(julianday(ship_date) -

Output:

Showing 4 rows

quarter_number	avg_shipping_time
1	57.16774193548387
2	71.11068702290076
3	117.75545851528385
4	174.09547738693468

4.10.3. Observations:

- Lowest average shipping time taken (57.16 days) in 1st quarter where as highest being in the last quarter (175.09 days).

5. Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
48610993.78	1000	994	3.135
Last Quarter Revenue	Last quarter Orders	Average Days to Ship	% Good Feedback
8573149.28	199	97.964	21.5

Table1: Business Metrics

Observations:

- Total revenue is 48610993.78.
- Total numbers of order are 1000.
- Total number of customers is 133.
- Average rating out of 5 is 3.13.
- Last quarter revenue is 8573149.28.
- Number of last quarter order is 199.
- Average time to ship is 97.96 days.
- Percentile good feedback is 21.5.

6. Important Highlights

State having highest customers	Top Maker as per no of order	Top selling vehicle in top selling state	Quarter with Highest AVG rating	Quarter with Highest positive customer feedback	Highest number of order	Number of orders
California	Chevrolet	Pontiac	1st quarter	1st quarter	1st quarter	310
Highest impact on revenue drop	Least impact on revenue drop	Highest revenue	Lowest revenue	Credit card with best discount	Highest shipping time	Lowest shipping time
3rd quarter	4th quarter	18032549.9	8573149.28	instapayment	175.09	57.16

Table2: Highlights

7. Business Recommendations

Improve After-Sales Service

- Enhance service quality by offering extended warranties and partnering with service centers. This will boost customer satisfaction and retention.

Address Negative Customer Feedback

- Analyze negative reviews each quarter. Use quick follow-ups with dissatisfied customers to increase ratings and improve customer perception.

Introduce Targeted Promotions

- Launch promotional campaigns for Q3 and Q4 to counteract revenue dips. Seasonal discounts and loyalty programs may drive more orders in slower quarters.

Focus on Popular Brands and Regions

- Prioritize popular brands based on the region. This focus will cater to customer preferences and maximize sales.

Adjust Discount Strategies by Payment Method

- Target discount offerings by payment type. This could incentivize more purchases and increase customer loyalty.

Optimize Shipping for Consistency

- Reduce shipping times in later quarters by improving logistics. Consistent delivery times will enhance customer satisfaction.

These actions can help New-Wheels stabilize revenue, improve customer satisfaction, and enhance its competitive position in the resale vehicle market.